

AD-A225 246

REPORT DOCUMENTATION PAGE

Form Approved

OMB No 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE	3. REPORT TYPE AND DATES COVERED JOURNAL ARTICLE																					
4. TITLE AND SUBTITLE MODIFIED NEGLECT OF DIATOMIC OVERLAP-TYPE SEMIEMPIRICAL METHODS			5. FUNDING NUMBERS																					
6. AUTHOR(S) KENNETH M. DIETER JAMES J.P. STEWART																								
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Frank J. Seiler Research Laboratory USAF Academy, CO 80840-6528			8. PERFORMING ORGANIZATION REPORT NUMBER FJSRL-JR-90-0014																					
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) AF OFFICE OF SCIENTIFIC RESEARCH BOLLING AFB, DC 20332			10. SPONSORING / MONITORING AGENCY REPORT NUMBER																					
11. SUPPLEMENTARY NOTES																								
12a. DISTRIBUTION / AVAILABILITY STATEMENT DISTRIBUTION UNLIMITED			12b. DISTRIBUTION CODE																					
13. ABSTRACT (Maximum 200 words) The current status of the semiempirical methods pioneered by Michael J.S. Dewar is given. These methods are made available to non-theoreticians through the programs MOPAC and AMPAC. Some capabilities of MOPAC and the form of the data input to the program are outlined.																								
<table border="1"> <tr> <td colspan="2">Specialization For</td> </tr> <tr> <td>None</td> <td><input checked="" type="checkbox"/></td> </tr> <tr> <td>None</td> <td><input type="checkbox"/></td> </tr> <tr> <td>Unannounced</td> <td><input type="checkbox"/></td> </tr> <tr> <td>Justification</td> <td></td> </tr> <tr> <td colspan="2">By</td> </tr> <tr> <td colspan="2">Distribution/</td> </tr> <tr> <td colspan="2">Availability Codes</td> </tr> <tr> <td colspan="2">Avail and/or</td> </tr> <tr> <td>Dist</td> <td>Special</td> </tr> </table>					Specialization For		None	<input checked="" type="checkbox"/>	None	<input type="checkbox"/>	Unannounced	<input type="checkbox"/>	Justification		By		Distribution/		Availability Codes		Avail and/or		Dist	Special
Specialization For																								
None	<input checked="" type="checkbox"/>																							
None	<input type="checkbox"/>																							
Unannounced	<input type="checkbox"/>																							
Justification																								
By																								
Distribution/																								
Availability Codes																								
Avail and/or																								
Dist	Special																							
14. SUBJECT TERMS A-1 20			15. NUMBER OF PAGES 11																					
			16. PRICE CODE																					
17. SECURITY CLASSIFICATION OF REPORT UNCLASSIFIED	18. SECURITY CLASSIFICATION OF THIS PAGE UNCLASSIFIED	19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED	20. LIMITATION OF ABSTRACT																					

"reaction coordinate", an approximation to the reaction profile is obtained, from which the heat of activation can be inferred.

Transition State Optimization. While the approximate geometry of the transition state (TS) can be obtained from the geometry corresponding to the maximum in a reaction path calculation, the actual reaction coordinate will normally involve more than one internal coordinate. Consequently, further optimization is normally required to determine the actual TS geometry. This geometry will correspond to the point of minimum energy for 3N-5 coordinates and a maximum in energy in the one coordinate leading from the reactant(s) to the product(s). The gradient norm for the system is still zero but the total energy is not at a minimum. Normal geometry optimization will not work because that procedure reduces both the gradient norm and the energy. In other words, using that procedure would result in the reoptimization of the reactants or products, depending upon which side of the actual TS the approximate TS geometry lies. To solve this problem, different procedures are available to optimize the TS geometry by minimizing only the gradient norm. The difference in the ΔH_f^\ddagger of the reactants and the TS is, then, the heat of activation predicted by the SE-MO model.

Force Constant Calculation. The characterization of specific points on the potential energy surface (PES) is in doubt until a force constant calculation is accomplished. The optimized geometry of a stable molecule corresponds to a local minimum on a PES and will, consequently, have all positive force constants. Transition states, on the other hand, will have a single negative force constant corresponding to the movement of the atoms along the reaction coordinate, since energy is at a maximum along that coordinate. It is also possible to "optimize" to a geometry which will have more than one negative force constant. These geometries normally correspond to points of no chemical significance. Consequently, procedures for force constant calculations are provided to allow users to characterize critical points along a reaction profile to ensure proper geometries, and corresponding values of properties of interest, have been calculated. In addition, vibrational frequencies are calculated from the mass weighted force matrix for comparison with experimentally observed spectra.

Carrying out a MOPAC Calculation

The data set. The data requirements for a MOPAC job are highly standardized. The calculation is controlled mainly by keywords. For example, if an unrestricted Hartree-Fock calculation is to be done on an ammonium ion using the AM1 method, then the keywords UHF, CHARGE=1, and AM1 would be used.

Two lines are provided to allow documentation of the system. Typically, this would consist of the chemical name of the system and the reason why the calculation was being done.

This is followed by the geometry, in either cartesian or internal coordinates. Each atom in the system is entered on a separate line. If internal coordinates are used, the atom's position is defined relative to other atoms in terms of a distance and two angles. If Cartesian coordinates are used, the position of each atom is defined relative to some arbitrary origin in Cartesian space.

Construction of the coordinate matrix, or Z-matrix, is the most difficult step in carrying out a MOPAC calculation. To make this task easier, special routines and programs have been written. One of these is DRAW.

Computational Requirements. With minor modifications, these programs all run on computers which support FORTRAN-77. With 1 megabyte (Mb) of storage very small molecules can be run using MOPAC. For molecules with 20 – 100 atoms, between 1.6 Mb and 16 Mb of storage are required. Estimates of computer time vary widely depending on the computer used and the nature of the calculation, however an idea of the time needed can be given by an example of a calculation on a VAX 11-780. For cyclohexane, on a VAX 11-780, using C-C distances on 1.5 Å, C-H of 1.1 Å, all angles tetrahedral, and no symmetry used, the time required to optimize the geometry is about 15 minutes. In general, the time required will rise as the cube of the number of atoms.

DRAW. Like MOPAC, DRAW is public domain software. Copies can be obtained from the Quantum Chemistry Program Exchange (QCPE) at the University of Indiana. DRAW allows the data input and output to be represented graphically. Because of the difficulty in assembling a valid Z-matrix, DRAW has been provided with an editor which allows a user to see how the geometry of the system changes as the coordinates are changed.

Other Graphics Programs. Various other programs are commercially available, such as SYBYL from Tripos and CHEM-X from Chemical Design, which, in addition to allowing MOPAC or AMPAC data-files to be easily generated, have the capability of doing a molecular mechanics energy minimization. This ability is very useful for 'cleaning up' user-written geometries prior to carrying out a semiempirical calculation, thereby reducing the time required to produce fully optimized geometries.

Accuracy of Semiempirical Methods

A knowledge of the accuracy, strong points, and weak points, of each method is necessary in order to efficiently carry out computational chemistry research. We will first look at a summary of the three most accurate methods in MOPAC, and then at their strengths and weaknesses.

Summary of MNDO, AM1, and PM3

Heats of Formation. Average errors for representative systems are given in Table I. From this, we see that the ΔH_f of simple organic compounds are predicted on average with chemically useful accuracy. Experimental values of ΔH_f apply to gas phase species at 298K. In order to see how well these methods work in general, the average error in ΔH_f for a wide range of systems, representing the chemistries of H, C, N, O, F, Si, P, S, Cl, Br, and I, are given. These systems include representative anions, cations, and radicals.

Molecular Geometries. Both MM and ab initio methods are more accurate at predicting molecular geometries than semiempirical methods. However, as stated earlier, MM methods are limited in their scope, while ab initio methods are generally too expensive. On average, semiempirical methods can predict bond lengths within about 2%, and angles within 3°. Many angles are defined by symmetry; these are not included in the analysis.

Dipole moments and Ionization potentials. Semiempirical methods predict dipole moments within about 0.4 D of that observed experimentally. Ionization potentials are not so well predicted, errors averaging about 0.7 eV.

Strengths and Limitations of Specific Methods

Strengths of MNDO. Unlike its precedent, MINDO/3, MNDO (and the later methods) was based on atomic rather than diatomic parameters. This enabled it to be parameterized for a large number of elements. Additionally, MNDO was the first semiempirical method to be based on molecular rather than atomic data. This gave it a generality other methods lacked. Many properties, such as molecular geometry, dipole, polarizability and hyperpolarizability, activation barriers, vibrational frequencies, and first and higher ionization potentials could now be calculated using a single program and a single method.

Limitations of MNDO. From its inception, some important limitations of MNDO were apparent. Sterically crowded molecules were calculated too unstable; for example, the ΔH_f of neopentane is predicted by MNDO to be -24.6 kcal/mol, compared with the observed -40.3 kcal/mol. On the other hand, four-membered rings were predicted to be too stable, this reaching a limit in cubane, which was predicted to be 49.6 kcal/mol too stable. Later on, other limitations were discovered, the most important from a biochemical standpoint being the virtually complete lack of a hydrogen bond. Other deficiencies included the extreme instability of hypervalent molecules. This effectively precluded the application of MNDO to organophosphorus compounds of biologic interest. Finally, activation barriers were predicted to be too high.

Strengths of AM1. AM1 is the first semiempirical NDDO method to accurately reproduce the heat of dimerization of water. This was achieved by modifying the core-core interaction. This change also corrected the instability of sterically crowded molecules, and the excessive stability of four membered rings. Average errors in ΔH_f obtained using AM1 are reduced by almost 40%, relative to MNDO.

Limitations of AM1. AM1 is still a relatively new method. As such, only a few limitations have become apparent thus far. One of the more important of these are the continued inability to adequately represent hypervalent chemistry. Very recently, AM1 parameters for phosphorus and sulfur have become available. Preliminary results indicate these are a considerable improvement over the MNDO method for predicting the ΔH_f of hypervalent compounds. Unfortunately, the phosphorus parameters include a barrier at 3.0\AA which complicates the interpretation of reaction mechanisms. For example, in the reaction $\text{PF}_3 + \text{F}_2 = \text{PF}_5$, there are two spurious barriers which must be crossed before the real barrier to the reaction is encountered. In some cases, for example in the reaction $\text{P}_2 + \text{P}_2 = \text{P}_4$, the spurious barrier is significantly higher than the true barrier.

Strengths of PM3. PM3 is the first NDDO method to adequately treat hypervalent systems. This is exemplified by sulfuric acid, for which the PM3 error in ΔH_f is -5.8 Kcal/mol. On the whole, PM3 is more accurate than

either AM1 or MNDO, the average error in ΔH_f being reduced by just over 40% relative to AM1, and 63% relative to MNDO.

Limitations of PM3. At the present time, PM3, the most recent NDDO method, has one severe limitation: the lack of adequate testing. At present only information about ground state systems is available.

Other limitations include the very different charge distribution compared to that given by MNDO and AM1. Some MM methods use semiempirical charge distributions, with various parameters being based on these charge distributions. Since the PM3 charge distribution is very different from that of MNDO or AM1, MM methods based on these charges will not be compatible with the PM3 charges.

Application to Carbohydrate Chemistry

The authors, not being familiar with carbohydrate chemistry, have limited this to a general discussion of two potential applications of semiempirical methods to the study of carbohydrate chemistry.

With the advent of PM3, biochemical reactions, for example, those involved in the Embden-Meyerhof pathway, can be studied. Until now, systems such as glucose-6-phosphate were either poorly represented, or were prohibitively slow to calculate.

As with most other computational methods, care must be exercised in the application of these techniques. Calculations assume isolated molecules, i.e. molecules in a vacuum, at absolute zero. Consequently, although the ΔH_f applies to the system at 298K, kinetic energy is not taken into account. However, calculated activation barriers can be used to predict relative reaction rates at 298K.

High polymer calculations can be performed on polysaccharides. Calculation of unit cell translation vectors (15), heats of polymerization (15), and elastic moduli (16) can readily be done. The accuracy of such calculations is the same as that of equivalent molecular species. A limitation of elastic moduli calculations is that the polymer is assumed to be 100% ordered, a state not commonly found in polysaccharides.

Summary.

At the present time, it appears that the applicability of semiempirical methods to the study of carbohydrate chemistry has been neglected. Methods are now available for the non-theoretician to investigate molecular systems, reaction mechanisms, and fundamental physical properties, without the need for any extensive knowledge of theoretical methods. Despite this, most computational studies appear to be limited to the use of molecular mechanics techniques. This is probably a natural consequence of the high accuracy and high speed of MM calculations. However, in recent years the ease of use, accuracy, and generality of semiempirical methods have improved considerably. While MM techniques should continue to be used for the study of ground state systems, carbohydrate chemists should be aware of the potential of semiempirical methods as a research tool, particularly for the study of reactions.

Literature Cited

1. For a recent review of these methods, see Clark, T. A Handbook of Computational Chemistry; John Wiley and Sons: New York, 1985.
2. For a more extensive review of semiempirical methods, see Thiel, W. Tetrahedron 1988, **44**, 7393.
3. For a good introduction to these techniques, see Pople, J. A.; Beveridge, D. L. Approximate Molecular Orbital Theory; McGraw-Hill: New York, 1970.
4. Bingham, R. C.; Dewar, M. J. S.; Lo, D. H. J. Am. Chem. Soc. 1975, **97**, 1285.
5. Dewar, M. J. S.; Thiel, W. J. Am. Chem. Soc. 1977, **99**, 4907.
6. Dewar, M. J. S.; Thiel, W. J. Am. Chem. Soc. 1977, **99**, 4899.
7. Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. J. Am. Chem. Soc. 1985, **107**, 3902.
8. Stewart, J. J. P. J. Comp. Chem. 1989, **10**, 209.
9. Stewart, J. J. P. J. Comp. Chem. 1989, **10**, 221.
10. Dewar, M. J. S. J. Phys. Chem. 1985, **89**, 2145.
11. Thiel, W., Program MNDO88, version 2.1, June 1988.
12. Available through QCPE, Department of Chemistry, Indiana University, Bloomington, IN 47405.
13. Available through QCPE.
14. Koopmans, T. Physica (Utrecht) 1934, **1**, 104.
15. Stewart, J. J. P. New Polymeric Materials 1987, **1**, 53.
16. Klei, H. E.; Stewart, J. J. P. Int. J. Quant. Chem. 1986, **20**, 529.

RECEIVED February 13, 1990

Reprinted from ACS Symposium Series No. 430
Computer Modeling of Carbohydrate Molecules
Alfred D. French and J. W. Brady, Editors
Published 1990 by the American Chemical Society

Chapter 3

"Modified Neglect of Diatomic Overlap"-type Semiempirical Methods

Kenneth M. Dieter and James J. P. Stewart

Frank J. Sellar Research Laboratory, U.S. Air Force Academy, CO 80840

The current status of the semiempirical methods pioneered by Michael J. S. Dewar is given. These methods are made available to non-theoreticians through the programs MOPAC and AMPAC. Some capabilities of MOPAC and the form of the data input to the program are outlined.

Chemists seeking to use computational chemistry to support experimental efforts now have three general theoretical tools available to them: force field or molecular mechanics models, *ab initio* molecular orbital (MO) models and semiempirical MO models (1). Each of these tools have strengths and weaknesses which must be evaluated to determine which is most appropriate for a given applications.

At one end of the spectrum are the molecular mechanics techniques, such as AMBER, CHARM, MODEL, and the MM programs, which use classical mechanics relationships to describe interactions between atoms in a chemical system. The associated algorithms include parameters which are optimized to reproduce experimental energies and geometries. These techniques are extremely fast and can accurately calculate energies of ground state systems. Heats of reaction and relative conformational stabilities can be derived for even very large systems such as enzymes. Without a quantum mechanical foundation, however, no information is available on electronic structure or properties dependent on electronic structure. Processes involving bond making/breaking cannot be represented, so full reaction profiles cannot be modeled.

At the other end of the spectrum are the *ab initio* quantum molecular models, which are rigorous within the Hartree-Fock/ Roothaan-Hall (HF/RH) formalisms. Electronic structure is calculated, and dependent properties are derivable. In theory, full reaction profiles can be modeled. In practice, however, their speed makes it impractical to apply the more accurate

This chapter not subject to U.S. copyright
Published 1990 American Chemical Society

90 05 14 052

methods to systems larger than, for example, acetic acid. The computing time required for these calculations is on the order of thousands to hundreds of thousands times that required for molecular mechanics calculations, depending on the basis set chosen for the ab initio calculations. There is always the temptation to use a smaller, less adequate basis set or to make simplifying assumptions about geometries or symmetry during a reaction to speed the calculations. Without proper justification, however, these procedures bias the calculations and cast doubt on the conclusions drawn from the study.

Semiempirical MO (SE-MO) calculations lie between these two extremes. Like ab initio models, SE-MO models are quantum mechanical in formalism. Like molecular mechanics, algorithms include parameters optimized to reproduce experimental values of molecular properties. The quantum mechanical foundation enables the calculation of electronic structure and derived properties, while, at the same time, the use of approximate functions for the interaction integrals reduces computing time by three or more orders of magnitude as compared to comparable ab initio calculations. While the computing time is still far too great to allow treatment of systems the size molecular mechanics can model, it is practical to look at real-life systems as opposed to being forced to modeling them with very small molecules. The concern is, of course, that the loss of rigor in the development of the SE-MO models will make them unrealistic. The forms of the parametric functions coupled with the optimization of parameters to reproduce experimental results is intended to compensate for these approximations. Additionally, the calculations are sufficiently inexpensive as to allow for extensive testing of these models to determine the validity of their application to specific studies.

Given, then, that SE-MO models are the only ones sufficiently fast to allow for the study of carbohydrate systems, while at the same time being able to provide information on electronic structure and reaction profiles, the remainder of this chapter will deal with the most commonly used models.

Historical Development

The history of general purpose SE-MO models is virtually totally the history of the work of Dewar and coworkers (2) based on the Intermediate Neglect of Differential Overlap (INDO) (3) and the Neglect of Diatomic Differential Overlap (NDDO) (3) approximations. While Pople and coworkers first developed these approximations to the full HF/RH treatment in the mid-sixties, it remained until 1975 before a model capable of calculating a variety of properties for a wide range of compounds became available. Dewar called this model MINDO/3 (4), for Modified INDO/version 3. It was eventually parametrized for 10 elements, but not in all combinations. A survey of molecules (5) containing C, H, N and O showed that the average error in heats of formation, ΔH_f , was 11 kcal/mol, in bond lengths, 0.022 Å, and in ionization potential, 0.7 eV.

MINDO. Despite its success, Dewar recognized certain weaknesses (6) in MINDO/3 due to the INDO approximation, such as the inability to model lone pair - lone pair interactions. Additionally, due to the use of diatomic parameters in MINDO/3, it was increasingly difficult to extend MINDO/3 to additional elements. Because of this, Dewar began working on a new model based on the better NDDO approximation.

This new model (6), called MNDO for Modified Neglect of Diatomic Overlap, was published by Dewar and Thiel in 1977. With MNDO the average errors (5) for the same survey of C, H, N and O molecules decreased to 6.3 kcal/mol for ΔH_f , 0.014 Å for bond lengths and 0.48 eV for ionization potentials. Since MNDO used only atomic parameters, parameterization of MNDO to include additional elements was much easier than with MINDO/3, and, over the next eight years, parameters were optimized for 16 elements in addition to C, H, N and O.

AM1. While MNDO was widely accepted and extensively used, there were still some deficiencies in the model. In particular, excessive repulsions were observed in MNDO potential energy surfaces just outside chemical bonding distances. This deficiency manifested itself (5,7) in the inability of MNDO to model hydrogen bonding, as well as in large positive errors in the ΔH_f of sterically crowded molecules and in heats of activation. Again Dewar set off to correct this deficiency.

The result of this effort was AM1, for Austin Model 1, published in 1985 (7). This model was essentially a reparameterized version of MNDO with a small change in the core-core repulsion function designed to eliminate the spurious repulsions. With this change and the more effective parameterization possible with a better optimization procedure, average errors in calculated properties again decreased. More importantly, AM1 was able to reproduce hydrogen bonding, making possible for the first time the study of biochemical reactions using these methods. Subsequently, AM1 has been, and still is being, expanded to include additional elements.

PM3. Throughout this work, parameters were optimized for a few elements at a time while fixing all previously optimized parameters. This was largely due to the combination of computer resources and optimization procedures available. Recently a new optimization algorithm was developed which allowed the simultaneous optimization of parameters for a larger number of elements (8). Using this procedure, a new model was parameterized for 12 elements (9). This model is called MNDO-PM3 to indicate it is the third parameterization of MNDO, AM1 being the second. In addition to improving performance for some classes of compounds, such as nitro compounds, which have presented problems for the other models, PM3 appears to model hypervalent compounds more accurately than ever before. The statistical distribution of errors in calculated ΔH_f is also significantly tighter than with the other models in a survey of over 500 compounds. While these results appear promising, PM3 will be validated only through extensive testing and application similar to that experienced by MINDO/3, MNDO and AM1.

Accuracy of NDDO-Based Models

Although both MNDO and AM1 have been used for several years, PM3 is still very new. A brief summary of the level of accuracy to be expected from calculations involving these methods is given in Table I. Note that average errors are given only for those elements for which parameters are available for all three methods.

Computer Programs. As Dewar has often stated (7,10), his whole intention in this work has been to develop a tool fast enough, simple enough and accurate enough for experimental chemists to use as an aid in their own research. A major contribution towards this goal in the last few years has been the combination of separate programs, written by Dewar and his coworkers to

Table I. Average errors in calculated properties

Heats of Formation (kcal/mol)				
Type of compounds	No. of compounds	Average errors		
		PM3	MNDO	AM1
Saturated hydrocarbons	24	4.5	7.0	6.7
Hydrocarbons with double bonds	42	2.8	4.1	4.4
Aromatic hydrocarbons	7	4.1	2.7	4.2
Alcohols	7	1.7	5.8	6.8
Ethers	8	4.0	10.1	5.8
Aldehydes and ketones	15	4.6	4.5	4.5
Acids	10	3.1	2.8	4.9
Compounds containing H, C, and O, only	57	4.5	5.6	6.0
Organophosphorus-V compounds	15	10.9	53.9	15.5
Compounds of C, H, N, and O, only	276	5.5	11.2	7.5
Wide range of compounds	713	8.2	22.5	13.8

Molecule Bond Lengths (Å)

Type of bond	No. of bond	Average errors		
		PM3	MNDO	AM1
C-H	51	0.009	0.010	0.014
C-C	72	0.017	0.014	0.017
C-O	18	0.012	0.021	0.030
O-H	8	0.014	0.014	0.012
O-P	3	0.026	0.045	0.044

Table I. Average errors in calculated properties (*Continued*)

Molecule Angles (Degrees)

Type	No.	Average errors		
		PM3	MNDO	AM1
H-C-H	4	1.2	4.8	4.1
H-C-C	19	1.3	1.7	1.3
H-C-O	3	3.3	2.7	2.9
C-C-C	16	2.3	4.1	2.4
C-C-O	4	2.1	0.7	1.0
C-O-H	2	2.9	7.3	2.3
C-C-C-C	7	8.2	22.6	8.8

Dipole Moments (D)

Type	No.	Average errors		
		PM3	MNDO	AM1
Compounds of C, H, N, and O, only	47	0.29	0.32	0.25
Wide range of compounds	132	0.37	0.44	0.35

Ionization Potentials (eV)

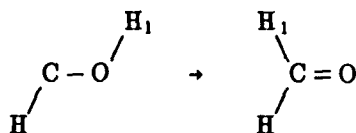
Type	No.	Average errors		
		PM3	MNDO	AM1
Compounds of C, H, N, and O, only	101	0.58	0.64	0.48
Wide range of compounds	264	0.56	0.77	0.59

implement these models, into comprehensive, user-friendly programs. There are currently three such programs: MNDO88, developed by Thiel at Wuppertal in West Germany (11); AMPAC (12), under continuing development by Dewar; and MOPAC (13), under continuing development in our laboratory after having its genesis in the Dewar group. The following are some of the major capabilities available within MOPAC and AMPAC. MNDO88 has similar capabilities.

Capabilities of MOPAC

Geometry Optimization. The most commonly used function of SE-MO calculations is the optimization of molecular geometries and the calculation of properties, such as ΔH_f , corresponding to the optimized geometries. The calculation starts with an approximate geometry input by the user using either cartesian coordinates or, more commonly, internal coordinates (bond lengths, bond angles and dihedral angles) to define the geometric points at which atoms are located. The forces acting on the system are calculated and, using this information, the geometry is changed incrementally to reduce the total energy. The geometry is optimized when the energy can no longer be reduced, that is, when the norm of the first derivatives of energy with respect to geometric parameters (gradient norm) is zero. In reality it is extremely difficult to reduce the gradient norm to zero. Additionally, because of the limitations of the quantum mechanical algorithms and the optimization algorithm, as well as the limited precision of the computer, the minima of the gradient norm and the energy will not necessarily coincide. In practice, the difference in geometries is very small; as a result, the gradient norm at the energy minimum will also be small. However, in some cases this can result in an optimization taking an excessively long time. The programs monitor the changes and anticipated changes in the gradient norm and energies and stop the calculations when they fall below preset limits. These limits can be adjusted to loosen or tighten the criteria as required for specific studies. The program then outputs the optimized geometry, ΔH_f , ionization energy based on Koopmans' theorem (14), charge distribution and dipole. Many other properties, such as delocalized and localized MOs, bond orders, etc., can also be output at the user's request.

Reaction Paths. A natural extension to the optimization of geometries and calculation of corresponding properties is the calculation of a reaction profile. A particular geometric coordinate approximating the reaction coordinate is changed incrementally to observe the corresponding change in system energy. For example, in the conversion of hydroxymethylene to formaldehyde, the O-C-H₁ bond angle can be fixed at various values ranging from its value in the optimized geometry of hydroxymethylene to that in the optimized geometry of formaldehyde while optimizing all other geometric coordinates.



When the corresponding energies for these points are plotted against the